UDP-GlcNAc Analogues as Inhibitors of O-GlcNAc Transferase (OGT)

Ghirardello, Mattia; Perrone, Daniela; Chinaglia, Nicola; Sádaba, David; Delso, Ignacio; Tejero, Tomas

Published in:
Chemistry: a European Journal

DOI:
10.1002/chem.201801083

Publication date:
2018

Document Version
Peer reviewed version

UDP-GlcNAc Analogs as Inhibitors of *O*-GlcNAc Transferase (OGT): Spectroscopic, Computational and Biological Studies

Mattia Ghirardello,[a] Daniela Perrone,[b] Nicola Chinaglia,[b] David Sádaba,[a] Ignacio Delso,[a] Tomas Tejero,[a] Elena Marchesi,[b] Marco Fogagnolo,[b] Karim Rafie,[c] Daan M. F. van Aalten[c] and Pedro Merino*[d]

Abstract: A series of glycomimetics of UDP-GlcNAc in which the β-phosphate has been replaced by either an alkyl chain or a triazolyl ring and the sugar moiety has been replaced by a pyrrolidine ring have been synthesized by using different click-chemistry procedures. Their affinity for human O-GlcNAc transferase (hOGT) has been evaluated and both spectroscopically and computationally studied. The binding epitopes of the best ligands have been determined in solution using saturation transfer difference (STD) NMR spectroscopy. Experimental, spectroscopic and computational results are in agreement, pointing out the essential role for binding of the β-phosphate. We have found that the loss of interactions from the β-phosphate can be counterbalanced by the presence of hydrophobic groups at a pyrroline ring acting as a surrogate of the carbohydrate unit. Two of the glycomimetics prepared reach inhibition in the micromolar scale.

Introduction

Protein-glycosylation, carried out by glycosyltransferases (GTs),[1] is a key post-translational modification in mammals.[2] Activation of the carbohydrate being transferred to the protein is produced by binding to a nucleotide diphosphate that interacts with the donor active site of the enzyme.[3] In the case of protein-glycosylation only nine nucleotide sugar donors (known as Leloir donors) containing uridine or guanine (with the exception of CMP-sialic acid, substrate for sialyltransferases) are substrates for glycosyltransferases.[4] Numerous pathological states including cancer[5] and neurological disorders[6] are closely related with the biosynthesis of glycoproteins. Inhibiting this essential biological process provides an obvious therapeutic target against these diseases.[7] Currently, the design and synthesis of inhibitors of glycosyltransferases is a hot topic of great interest.[8] As illustrated in Figure 1, several strategies are used to inhibit protein glycosylation.[9] The most used approach consists of modifying the carbohydrate unit (**1**).[10] Modifications at the ribose ring (**2**) and the nucleobase (**3**) of the nucleotide have also been reported but to a lesser extent.[11]

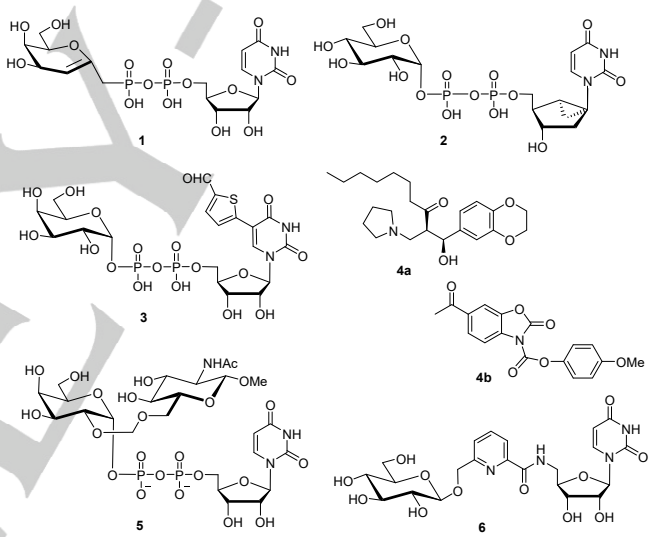


Figure 1. Strategies towards design of GT inhibitors.

More recently, new chemotypes like **4** far from GT substrate-like compounds[12] and bisubstrate inhibitors such as **5**,[13] consisting of two conjugated fragments interacting with both donor and acceptor sites of the enzyme, have emerged as a promising alternatives (Figure 1). Modifications at the pyrophosphate linkage is also possible. Indeed, it is known that pyrophosphate derivatives have poor membrane permeability and, consequently, low bioavailability. For that reason, conjugated sugar-nucleotides linked by neutral surrogates of the pyrophosphate unit like **6** have also been designed as GT inhibitors.[12b, 14]

In this context, we have studied[15] less-polar UDP-Gal and UDP-GalNAc analogs in which the β-phosphate has been replaced by an alkyl chain. As a model we chose GT GalNAc-T2, an important GT widely distributed in human tissues that play an important role

[a] Dr. M. Ghirardello, Dr. D. Sadaba, Dr. I. Delso, Prof. T. Tejero
Instituto de Síntesis Química y Catálisis Homogénea (ISQCH). Universidad de Zaragoza. CSIC. E-50009 Zaragoza, Spain.
[b] Dr. D. Perrone, Mr. N. Chinaglia, Dr. E. Marchesi, Dr. M. Fogagnolo
Department of Chemical and Pharmaceutical Sciences, Università degli Studi di Ferrara, I-44121 Ferrara, Italy
[c] Dr. K. Rafie, Prof. D. M. F. van Aalten
Centre for Gene Regulation and Expression, School of Life Sciences, University of Dundee, Dundee DD1 5EH, U.K
[d] Prof. P. Merino
Instituto de Biocomputación y Física de Sistemas Complejos (BIFI). Universidad de Zaragoza. E-50009 Zaragoza, Spain.
E-mail: pmerino@unizar.es

Supporting information for this article is given via a link at the end of the document.

in health and disease.[16] The study showed the analogs to be weaker binders of the enzyme than the natural substrate UDP-GalNAc and from the crystal structure it was inferred that the β-phosphate was required for binding to the metal ion.

Herein we use as the *O*-GlcNAc transferase (OGT) as a model system, an enzyme that plays a crucial role in a variety of biological functions within the human body[17] and does not require the presence of a metal.[18] Development of new inhibitors of OGT has been pursued for several years.[19] However, although promising results have been reported recently [20] finding specific and cell-permeable inhibitors to be used in animal models still remains a challenge.[21]

In the present study, we used a total of 14 candidates (Figure 2). UDP-Glc and UDP-GlcNAc analogs **7** and **8**, respectively, were chosen by analogy with our previous study with GalNAc-T2.[15] Some previously reported inhibitors consisted of replacing the sugar moiety by a different structure since it is known that the sugar unit has some conformational freedom.[22] On the other hand, it has been recently reported that the precise conformation of the sugar unit could play a significant role in catalysis.[23] To study the effects of sugar substitution we included compounds **9-16** in which the sugar moiety has been replaced by a pyrrolidine unit bearing functional groups adequate for H-bond (as in **9**, **11**, **13** and **15**) and hydrophobic interactions (as in **10**, **12**, **14** and **16**). Finally, analogs **17-20** containing a triazole ring in place of the β-phosphate have also been prepared (Figure 2).

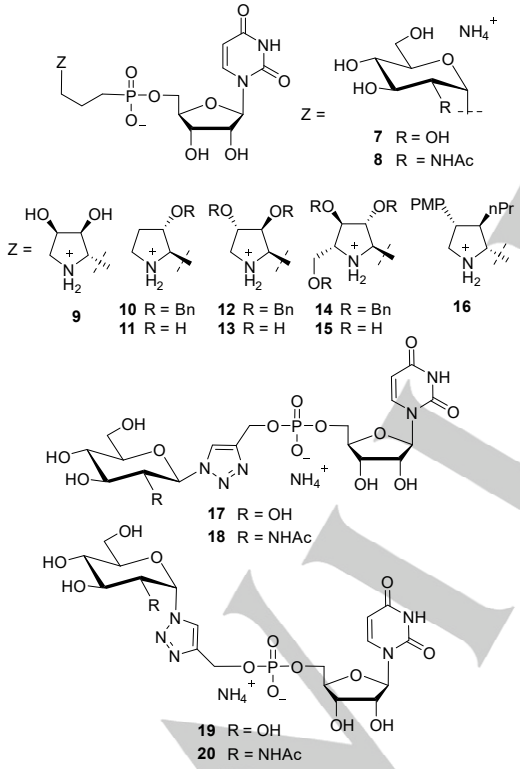


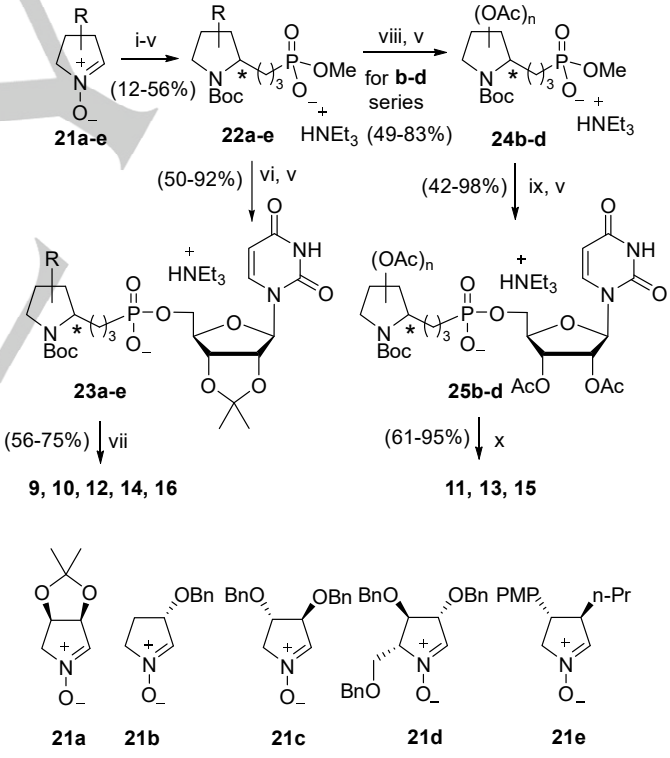
Figure 2. Less-polar analogs of UDP-Glc and UDP-GlcNAc.

Replacement of both phosphate groups by a triazole had been reported previously but the heterocyclic ring was placed close to the nucleoside and no biological studies were carried out.[24] More recently, Vidal and co-workers reported the synthesis of neutral analogs containing glycosyltriazoles, in which both phosphate groups were replaced and suggested that the presence of the triazole ring results in additional interactions when a divalent metal ion is required for catalysis[14a,20a]. Our approach also consists of employing glycosyltriazoles, however only the β-phosphate is replaced in order to determine the exact equivalence between the β-phosphate and the triazole ring in enzymes that do not require a metal ion for catalysis. The chemical synthesis, biological evaluation and computational studies of the prepared compounds **7-20** will be discussed in the present work.

Results and Discussion

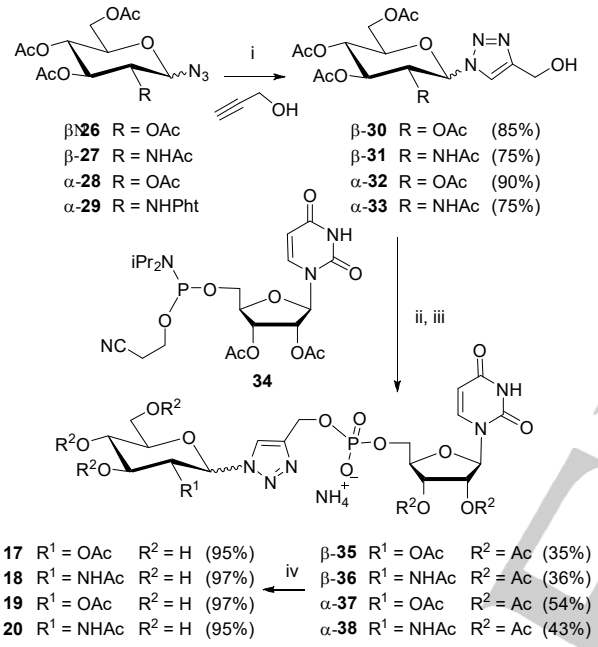
Synthesis of UDP-GlcNAc Analogs.

The synthesis of compounds **7** and **8** have been reported in our previous study with GalNAc-T2.[15] Novel compounds **9-16** were synthesized from nitrones **21a-e** through a highly diastereo-selective allylation[25,26] and photoinduced free-radical hydro-phosphonylation[27] as key steps (Scheme 1; for details see SI).



Scheme 1. *Reagents and conditions:* (i) allylmagnesium bromide, THF, -80 ºC, 2 h. (ii) Zn, AcOH, H_2O, 6 h, rt. (iii) Boc_2O, pyridine, DCM, 16 h, rt. (iv) dimethyl phosphite, DPAP, hv, 30 min, rt. (v) PhSH, Et_3N, dioxane, 4 h, rt. (vi) 2',3'-di-O-isopropylidene-uridine, BOP, DMF, iPr_2EtN, 4 h, rt. (vii) TFA, H_2O, 3 h, rt. (viii) H_2, $Pd(OH)_2$-C, Ac_2O, pyridine, 12 h, rt. (ix) 2',3'-di-O-acetyl-uridine, BOP, DMF, iPr_2EtN, 4 h, rt. (x) 25% NH_4OH, MeOH, 6 h , rt.

Selectively monodeprotected phosphonates **22a-e** were coupled with partially protected uridine to give advanced intermediates **23a-e**. Deprotection of the phosphate moiety and acidic treatment afforded analogs **9**, **10**, **12**, **14** and **16** in good chemical yields (see SI). The synthesis of completely deprotected analogs was carried out starting from hydrophosphonylated intermediates **22b-d**. Exchange of benzyl groups by acetyl groups was carried out through hydrogenation in the presence of acetic anhydride, compounds **24b-d** being obtained in excellent yields after partial deprotection of the phosphonate function. Coupling with partially protected uridine afforded intermediates **25b-d**. Complete deprotection of those intermediates yielded deprotected analogs **11**, **13**, and **15** in good chemical yields (see SI).



Scheme 2. *Reagents and conditions:* (i) for β-**26** and β-**27**: $(EtO)_3P{\cdot}CuI$, 1:1 toluene/THF, 24 h, reflux; for α-**28**: $CuSO_4$ (0.2 eq), sodium ascorbate (0.4 eq), THF, H_2O, 20-25 min, reflux; for α-**29**: $CuSO_4$ (0.2 eq), sodium ascorbate (0.4 eq), tBuOH, H_2O, 20 h, 50 ºC; then DMTrCl, pyridine, 18 h, rt, then $MeNH_2$ aq, 35%, MeOH, 18 h, 70 ºC, then Ac_2O, pyridine, 20 h, rt, then AcOH aq 80%, MeOH, 1 h, 50 ºC. (ii) DCI, MeCN, 45 min, rt. (iii) tBuOOH, decane, MeCN, 45 min, rt.; then Et_2NH, MeCN, 45 min, rt. (iv) NH_4OH, MeCN, 18 h, rt. DCI: 4,5-dicyanoimdazole

The synthesis of compounds **17-20** started from azides **26-29**, which were obtained from the corresponding carbohydrates (see SI). Intermediates **30-33** were obtained through a typical CuAAC "click" reaction[28] which has been widely employed in carbohydrate chemistry.[29] It was necessary to modify the reaction conditions slightly for each substrate. Whereas the classical copper(II) sulphate/sodium ascorbate system performed well for α-anomers, the reactions leading to β-anomers showed better results with the system $(EtO)_3P{\cdot}CuI$. In the case of compound α-**29** additional transformation of the phtalimido group into the acetamido group was necessary (Scheme 2). Coupling between glycosyltriazoles **30-33** and the nucleotide unit **34** was realized using the phosphoramidite methodology.[30] The resulting intermediates **35-38** were further transformed into the target analogues **17-29**.

Biological Assays (K_i measurements with human OGT in vitro).
Compound **7-20** were subjected to enzyme binding studies to the human O-GlcNAc transferase via fluorescence (Table 1).

Table 1. Enzymatic inhibition of human O-GlcNAc transferase (hOGT).

inhibitor	K_i (μM)[a]
7	920.7±58
8	4769.0±905
9	5452.0±1875
10	718.0±33
11	2218.0±595
12	1140.0±473
13	1804.0±110
14	102.5±19
15	1814.0±132
16	2262.4±10
17	1439.5±249
18	231.9±4
19	1048.0±475
20	589.1±63

[a] Apparent K_i values were calculated by introducing an 100% displacement at a fictive concentration of 10 M..

Binding affinities were determined for all the compounds by displacement of a fluorescent probe[31] by compounds **7-20**. The non-fluorescein modified probe was used as a positive control to show proper displacement. Most compounds showed no inhibition and only compounds **14** and **18** showed moderate binding. Due to the low binding affinity of the compounds, 100% displacement was mimicked by introducing an artificial concentration of 10 M, which allowed the fitting of a 4-parameter non-linear-regression curve fit and extrapolating apparent K_i values. Among the sugar analogs studied (**7**, **8**, **17-20**) the best result was found for the UDP-GlcNAc analogue **18** presenting a beta configuration at the anomeric center while the analogue **20** with alpha configuration (the same as in the natural substrate) was a poor inhibitor. In fact, analogues **7** and **8**, both with alpha configuration, showed rather high K_i values. Another noteworthy finding is the high K_i value for compound **8** for which the only difference with the natural substrate is the lack of β-phosphate. This finding clearly illustrates the key role of that phosphate unit in binding.

Pyrrolidinyl analogues showed high K_i values with the exception of compound **14** which showed a reasonable value of K_i = 102.5 μM, the tightest binder for all the UDP-GlcNAc analogues evaluated (the most potent inhibitor reported to date is the product of the reaction, UDP, showing K_d = 0.5 μM[32]). Presumably, the presence of benzyl groups might promote hydrophobic interactions that stabilize the ligand. Indeed, successive elimination of benzyl groups in compounds **12** and **10** resulted in a notable increase of K_i values. In the case of compound **12** the opposite configurations of the stereogenic centers of the pyrrolidine ring with respect to **10** and **14** might also affect the binding negatively. The high values observed for analogues **11, 13** and **15** suggest that H-bond interactions are not essential in that section of the binding pocket. Attempts to crystallise compound **14** in complex with hOGT revealed strong density for the UMP part in the binding site, however lacked the density for the rest of the ligand. Likely resulting from a high flexibility of the latter leading to a lack of electron density and therefore does not allow model building. Nevertheless, given the correct orientation of UMP it is possible to infer the orientation of the benzyl groups toward a hydrophobic pocket as computational calculations support (see below).

Spectroscopic studies (STD-NMR experiments).

Saturation transfer difference (STD) NMR spectroscopy[33] has demonstrated to be a great utility in evaluating ligand-protein binding affinities.[34] In particular, STD NMR allows the identification of binding epitopes,[35] which reflect the distances from ligand protons to protons of binding site residues. STD-NMR experiments were carried out with selected compounds to cover a range of K_i values, from the best ligand **14** (K_i =102 μM) to **12** (K_i 1.14 mM). UMP was selected as a reference instead UDP since the low K_d (0.5 μM) of the latter will prevent the observation of any displacement by the added ligand. Moreover, since the studied ligands lack the β-phosphate, the use of UMP as a reference will provide a correct evaluation of the interactions of the glycomimetic moiety. To obtain the epitope mapping of the selected ligands, UMP, **10**, **14** and **18** were incubated with hOGT at pH = 7.4 and the STD-NMR spectra were acquired at 25 ºC. Due to the low affinity, the experiments were carried out with a 1:50 enzyme/ligand ratio which yielded STD spectra with the best quality. The analysis of the STD intensities allowed the mapping of the epitopes of UMP and compounds **10**, **14** and **18** (Figure 3). The uridine moiety shows the same trend in STD intensities with 100% relative STD enhancement for H-1 in all cases. The best results with strong interactions for all the uridine protons were found for the best ligand **14**. On the other hand, compound **18** only showed interactions for H-1, H-2 and H-1'. Remarkably, compound **14** showed better recognition of the uridine moiety than UMP. Indeed, when **14** was added to a solution containing the complex UMP-hOGT, only signals resembling a hOGT **14** complex were observed in a higher intensity, showing that **14** displaces UMP and binds hOGT with higher affinity. These results agree with the crystallographic structure of the complex UDP-GlcNAc/hOGT,[17c] where the uridine hydrogens pointing towards the inner part of the protein show the highest intensities. The highest STD signals for the glycomimetic unit was also observed with **14** and correspond to the three phenyl rings suggesting the presence of important hydrophobic interactions (Figure 4, for the whole set of spectra see SI).

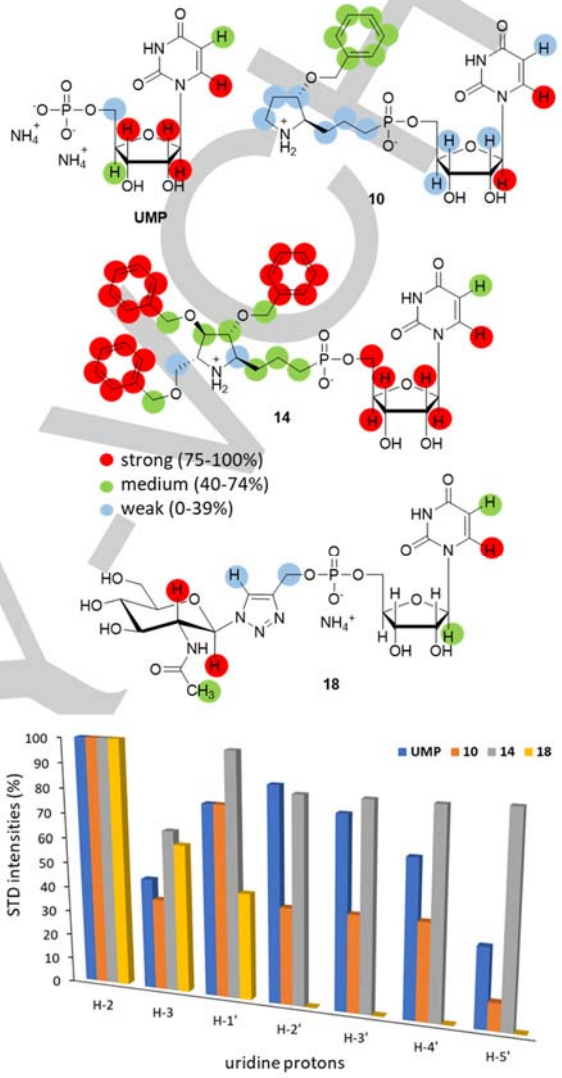


Figure 3. Top: Epitope mapping of UMP and compounds **10**, **14** and **18**. Bottom: Comparison of STD intensities (%) of the protons of uridine moiety.

In fact, elimination of two benzyl groups as in **10** resulted in a considerably loss of affinity also indicating that the remaining benzyl group at C-3 of the pyrrolidine ring was not essential in the binding. In the case of UDP-GalNAc analogue **18** only H-1 and H-2 of the galactosamine unit showed a strong interaction with the protein. The triazole ring did not show any remarkable interaction indicating that it does not occupy a favorable position for binding. No strong interactions were observed for the alkyl chain in place of the β-phosphate confirming the loss of binding due to elimination of that phosphate unit. Nevertheless, whereas a weak interaction was observed for **10**, limited interaction was observed in the case of **14**. STD data confirmed that a pyrrolidine ring with

the appropriate configuration, as seen in **14,** is an adequate scaffold for engaging the aromatic residues. However, other systems would be also suitable for the same task. These experiments suggest the presence of a hydrophobic pocket on hOGT and the importance of establishing hydrophobic interactions with the surrounding area of the sugar residue binding site.

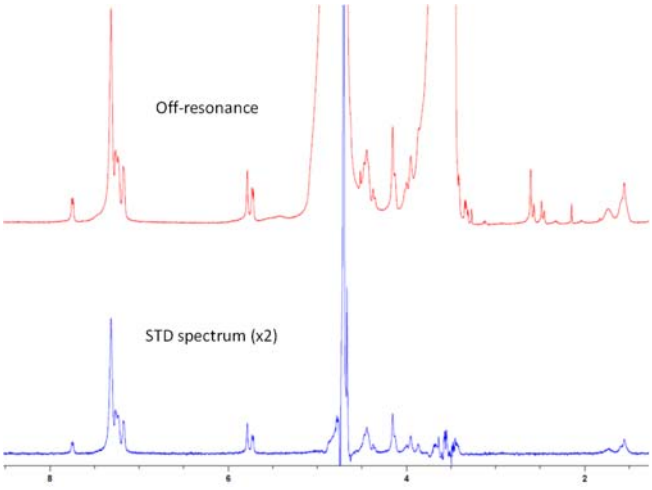


Figure 4. Off-resonance and STD NMR spectra of **14** in complex with hOGT.

Computational Studies (Docking and Molecular Dynamics).

To shed additional light on the protein-ligand binding modes, computational studies with the best ligands **14** and **18** were carried out using the known structure of OGT in complex with UDP-GlcNAc (PDB ID: 4GZ5)[17c] as the template (comparative studies were also carried out with UDP-GlcNAc, **10** and **12**; see SI). Published crystal structures of UDP-GlcNAc and UDP in complex with hOGT[17c] show interactions between the base moiety, and Ala896 (H-bond) and His901 (π,π-interactions). Previously reported docking studies using the crystal structure of OGT in complex with UDP-GlcNAc as a template, indicated the presence of four possible hydrogen-bonds between the GlcNAc residue and OGT, i.e: C6-OH/Thr560, C4-OH/Leu653, C3-OH/Gly654 and C2-N-acyl group/His920).[36] The same studies suggested that the backbone carbonyl oxygen of Leu653 and the hydroxyl group of Thr560 contribute to binding of UDP-GlcNAc through key H-bond interactions and that the C2-acetamido points to a hydrophobic pocket constituted by Met501, Leu502 and Tyr841. Docking studies with β-1-triazolyl-GlcNAc analogues suggested interactions between the triazole ring and Thr921 and Thr924 resembling those from β-phosphate of UDP-GlNAc.[20a] All these computational studies were based on docking calculations and were when we applied such a type of calculation using the Schrodinger 2017 software package. Docking with **14** and **18** was performed by including both ligands in the UDP-GlcNAc binding site. In both cases, it was observed that the orientation of the nucleoside-α-phosphate unit resembles the orientation of the natural substrate and the same interactions are observed for both UDP-GlcNAc and other ligands (for details see SI).

Molecular dynamics simulations allow studying the evolution of the complex and offers more valuable and representative information than static and minimum-energy docking calculations. In this context, MD simulations were previously used for elucidating the catalytic mechanism of OGT[37] further supported by structural data.[17c] These studies pointed out the key role of α-phosphate and Asp554 as the catalytic bases, which was further confirmed as the most energetically favorable pathway.[38]

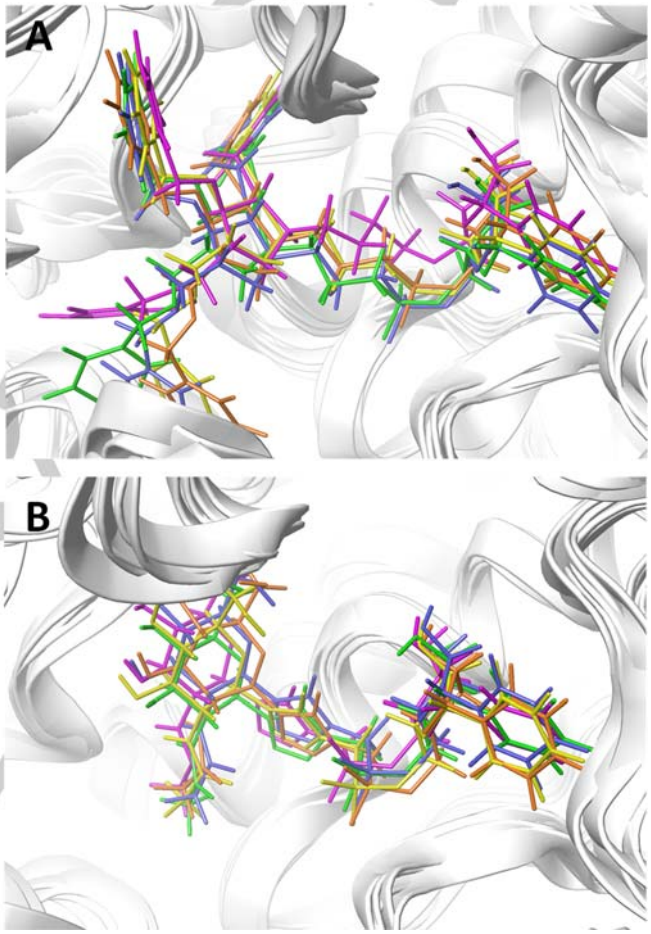


Figure 5. Superimposition of ten structures of OGT/**14** (**A**) and OGT/**18** (**B**) complexes along the MD simulation from 50 to 100 ns; intermediate structures are shown grading from orange (starting geometry), yellow, green, blue (intermediate geometries) to magenta (final target geometry).

On the other hand, when Asp554 was mutated out a complete loss of activity was not observed, resulting more plausible that α-phoshate could act as the catalytic base.[32] Recent structural studies carried out by one of us established that the peptide-binding site imposes size and conformational restrictions.[39] In our studies, we used the generated poses in the docking studies for running MD simulations for 100 ns. A preliminary analysis showed that, after 50 ns, the situation is stabilized and basically the same orientation is found for the ligands (Figure 5,A).

For compound **14**, however, different orientations were observed for the CH_2OBn group at C-5 of the pyrrolidine ring whereas the other benzyloxy groups at C-3 and C-4 remain in the same orientation. Compound **18** showed to be always in the same orientation (Figure 5, B). A comparison between **14** and **18** (Figure 6) revealed that both have the same preferred conformation for the mononucleotide unit (*exo* orientation of the nucleobase with respect to ribose unit with a dihedral angle C(C=O)-N(glycosydic)-C(anomeric)-O(ribose) around 180º). On the other hand, such a unit is oriented in a different way inside the binding pocket, although the same interactions were observed for this part of the molecule (see below). Notably, the glycomimetic unit in **8** (glycosyltriazol moiety) pointed towards the same direction than the less-fixed CH_2OBn group at C-5 of the pyrrolidine ring of **14**. This result agrees with the observed lower binding of **18** and supports the hypothesis that benzyl groups at C-3 and C-4 are required for binding, particularly the benzyl at C-4 as compound **12,** with only a single benzyl group at C-3 has a markedly lower affinity.

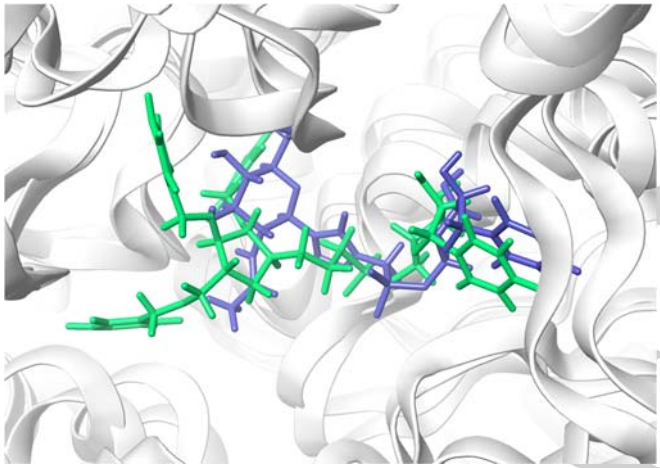


Figure 6. Superimposition of the OGT/**14** (green) and OGT/**18** (blue) complexes found at 100 ns.

A detailed inspection of interactions between OGT and **14** revealed the same interactions observed in the crystal structure for the mononucleotide unit (Figure 7, A). The nucleobase has H-bond interactions with Ala896 and Arg904, and hydrophobic interactions with His901. The ribose unit presents H-bond interactions with residues Lys898 (with O at C-2' and C-3') and Asp925 (with OH at C-3'). The phosphate group forms H-bond interactions with Lys842, Cys911, His920, Thr921 and Thr922. The three benzyl groups, responsible for the observed affinity are installed in hydrophobic pockets. An aromatic group at C-3 of the pyrrolidine ring is surrounded by Leu653, Phe694 and Thr560. The benzyl group at C-4 points towards a cavity formed by Ser553, Asp554, Thr560 and His558. The benzyloxymethyl group at C-5 is oriented towards the hydrophobic area formed by Pro497, His498, Met501, Thr633, Tyr655 and Pro656. Compound **18** (Figure 7, B) shows the same interactions for the nucleobase (Ala896, Arg904 and His901) and H-bond to Asp925 for the ribose unit.

On the other hand, interactions of the NHAc group of glucosamine moiety with His498 and His920, present in the crystal and MD simulations are lost in **18**, which form H-bond interactions with Tyr841 and Lys842. In contrast with previous docking studies[14a], however are in agreement with STD-NMR experiments, where no interactions were found for the triazole ring.

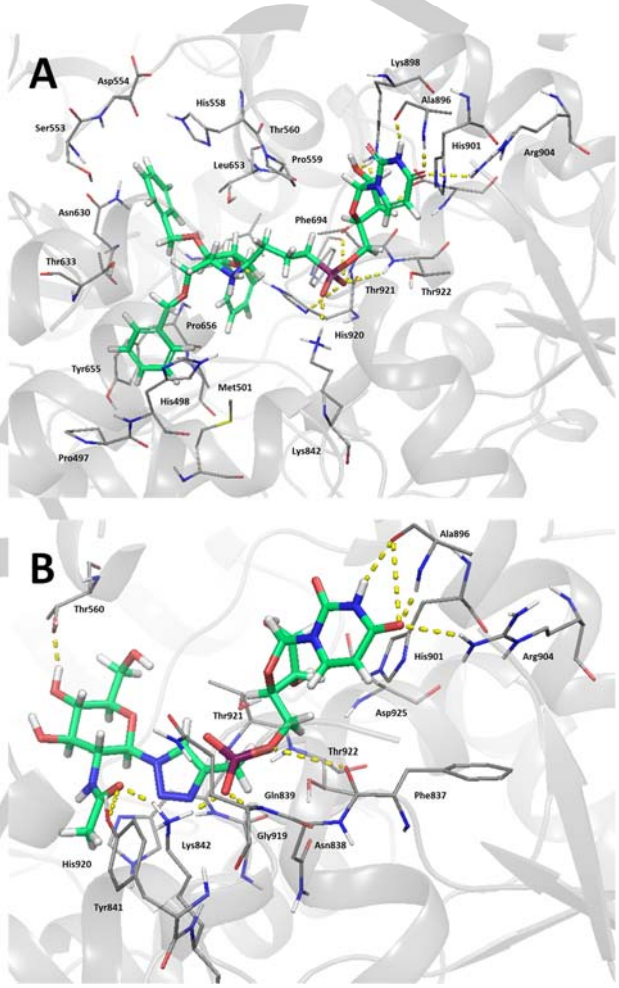


Figure 7. MD simulation for compound **14** (**A**) and 1**8** (**B**). Structures correspond to stable situations after 100 ns.

Conclusions

In the search of less-polar compounds to increase bioavailability of potential OGT inhibitors, we used click-chemistry procedures based on photoinduced free-radical hydrophosphonylation and copper catalyzed alkyne azide cycloaddition (CuAAC), to prepare a series of glycomimetics of UDP-GlcNAc in which the β-phosphate group has been replaced by an alkyl chain. It has been reported that both the acetamido group and the β-phosphate play key roles in binding hOGT whereas α-phosphate is crucial for catalysis.[32] The high K_i value found for compound **7** (OH instead NHAc) supports the key role of the acetamido group in the

binding.[17c] Indeed, it has been reported that OGT can transfer UDP-GalNAc but not UDP-Glc.[17c] However, the higher K_i value of **8**, which has the NHAc group suggests that it is the absence of the β-phosphate the responsible of a wrong orientation of the carbohydrate unit. The β-phosphate provides key hydrogen bonds (with Lys842 and three contiguous residues His920, Thr921, Thr922) that are observed in both the studied OGT-UDP-GlcNAc complex (PDB: 4GZ5) and the ternary complex OGT-UDP-GlcNAc-peptide (PDB: 4GYW).[17c] Both experimental and computational results confirm the crucial role of β-phosphate in binding for hOGT in a similar way to other glycosyltransferases requiring the presence of a metal ion, like GalNAc-T2.[15] Consequently, the elimination of the β-phosphate requires groups in the analogues capable of providing interactions to counterbalance those lost. In this respect, glycomimetics with free hydroxyl groups do not demonstrate to be good candidates suggesting that H-bond interactions are not the key in this case. On the other hand, the triazole ring give some interactions in the case of compound **18** (K_i = 231.9 μM) in agreement with that reported by Vidal and co-workers.[14a] However, the best result was found for compound **14** (K_i = 102.5 μM) with a value in the order to those reported in the literature.[14b] This result supported by STD NMR experiments and computational (docking and MD) calculations, points out the importance of hydrophobic interactions in the area of the active site occupied by the carbohydrate for the design of an inhibitor.

Our study is consistent with the lower affinity of compounds in which the β-phosphate unit has been replaced by a different moiety and reaffirm the importance of the interactions of such a unit even in the absence of a metal for achieving a good binding, as in the present case of hOGT for which it has been suggested that Lys842 adopts the role of the metal ion. At the same time the study shows the possibility of introducing hydrophobic groups at the carbohydrate surrogate to increase affinity. In other words, the lower affinity caused by the absence of β-phosphate is counterbalanced by the hydrophobic interactions of the poly-*O*-benzylated pyrrolidine acting as a surrogate of the carbohydrate unit. A combination of this approach with the replacement of the β-phosphate by a functional group which provide interactions capable of competing with those observed in the natural substrate might be the way to find inhibitors of hOGT at a nanomolar level.

Experimental Section

Data of Final Compounds **9-20** are given. For the complete experimental section see SI

Compound 9. Pale yellow oil. $[\alpha]_D^{29} \cong$ - 6 (*c* 1.3, H_2O). ^{1}H NMR (400 MHz, $D_2O^{25°C}$) *δ* 7.94 (d, 1H, *J* = 8.1 Hz, H-6_u); 5.95 (d, 1H, *J* = 4.8 Hz, H-1'); 5.93 (d, 1H, *J* = 8.1 Hz, H-5_u); 4.37-4.33 (m, 2H, H-2', H-4); 4.31 (t, 1H, *J* = 5.0 Hz, H-3'); 4.28-4.24 (m, 1H, H-4'); 4.15 (ddd, 1H, *J* = 11.8, 4.5, 2.5 Hz, H-5'); 4.09-4.04 (m, 1H, H-5'); 4.05 (dd, 1H, *J* = 8.6, 4.2 Hz, H-3); 3.52 (dd, 1H, *J* = 13.0, 4.4 Hz, H-5); 3.53-3.46 (m, 1H, H-3); 3.32 (dd, 1H, *J* = 13.0, 2.1 Hz, H-5); 1.99-1.62 (m, 6H, 2 H-6, 2 H-7, 2 H-8). ^{13}C NMR (100 MHz, $D_2O^{25°C}$) *δ* 166.2 (C); 151.7 (C); 141.6 (CH-6_u); 102.3 (CH-5_u); 88.9 (CH-1'); 83.1 (d, *J* = 7.7 Hz, CH-4'); 74.9 (CH-3); 73.8 (CH-2'); 69.5 (CH-3'); 29.2 (CH-4); 62.8 (d, *J* = 5.2 Hz, CH_2-5'); 60.2 (CH-2); 49.0 (CH_2-5); 30.8 (d, *J* = 16.4 Hz, CH_2); 25.4 (d, *J* = 134.9 Hz, CH_2); 20.1 (d, *J* = 4.4 Hz, CH_2). ^{31}P NMR (162 MHz, $D_2O^{25°C}$) *δ* 27.5. Anal Calcd for $C_{16}H_{26}N_3O_{10}P$: C, 42.58; H, 5.81; N, 9.31; P, 6.86. Found: C, 44.05; H, 5.84; N, 9.54; P, 6.69.

Compound 10. White solid. Mp: 162 °C, with decomposition. $[\alpha]_D^{30} \cong$ + 19 (*c* 1.4, H_2O). ^{1}H NMR (400 MHz, $D_2O^{25°C}$) *δ* 7.78 (d, 1H, *J* = 8.0 Hz, H-6_u); 7.36-7.23 (m, 5H, Ph); 5.87-5.68 (m, 1H, H-5_u); 5.79 (s, 1H, H-1'); 4.53-4.39 (m, 2H, C*H*$_2$Ph); 4.20-4.10 (m, 3H, H-2', H-3', H-4'); 4.07-3.99 (m, 2H, H-3, H-5'); 3.98-3.91 (m, 1H, H-5'); 3.60-3.52 (m, 1H, H-2); 3.39-3.24 (m, 2H, 2 H-5); 2.20-2.03 (m, 2H, 2 H-4); 1.72-1.47 (m, 6H, 2H-6, 2 H-7, 2 H-8); ^{13}C NMR (100 MHz, $D_2O^{25°C}$) *δ* 165.9 (C); 151.5 (C); 141.5 (CH-6_u); 136.7 (C); 128.8 (2 CH_{Ph}); 128.4 (CH_{Ph}); 128.3 (2 CH_{Ph}); 102.3 (CH-5_u); 89.1 (CH-1'); 83.0 (d, *J* = 7.7 Hz, CH-4'); 81.1 (CH-3); 73.8 (CH-2'); 71.2 (CH_2Ph); 69.3 (CH-3'); 64.1 (CH-2); 62.6 (d, *J* = 5.2 Hz, CH_2-5'); 43.6 (CH_2-5); 30.9 (d, *J* = 16.5 Hz, CH_2); 29.2 (CH_2-4); 25.4 (d, *J* = 135.1 Hz, CH_2); 20.1 (d, *J* = 4.3 Hz, CH_2). ^{31}P NMR (162 MHz, $D_2O^{25°C}$) *δ* 27.3. Anal Calcd for $C_{23}H_{32}N_3O_9P$: C, 52.57; H, 6.14; N, 8.00; P, 5.89. Found: C, 51.21; H, 6.25; N, 7.81; P,5.70.

Compound 11. White solid. mp: 171-172 °C. $[\alpha]_D^{28} \cong$ + 16 (*c* 1.1, H_2O). ^{1}H NMR (400 MHz, $D_2O^{25°C}$) *δ* 7.89 (d, 1H, *J* = 8.1 Hz, H-6_u); 5.90 (d, 1H, *J* = 5.0 Hz, H-1'); 5.88 (d, 1H, *J* = 8.1 Hz, H-5_u); 4.31 (t, 1H, *J* = 5.0 Hz, H-2'); 4.28-4.23 (m, 2H, H-3, H-3'); 4.23-4.18 (m, 1H, H-4'); 4.10 (ddd, 1H, *J* = 11.8, 4.1, 2.4 Hz, H-5'); 4.02 (ddd, 1H, *J* = 11.8, 5.5, 3.2 Hz, H-5'); 3.48-3.33 (m, 3H, H-2, 2 H-5); 2.21 (ddd, 1H, *J* = 14.0, 9.9, 5.7 Hz, H-4); 1.96 (ddd, 1H, *J* = 14.0, 10.3, 6.5 Hz, H-4); 1.83-1.54 (m, 6H, 2 H-6, 2 H-7, 2H-8). ^{13}C NMR (75 MHz, $D_2O^{25°C}$) *δ* 166.2 (C); 151.7 (C); 141.6 (CH-6_u); 102.4 (CH-5_u); 88.9 (CH-1'); 83.1 (d, *J* = 7.7 Hz, CH-4'); 73.8 (CH-2'); 73.6 (CH-3); 69.5 (CH-3'); 65.9 (CH-2); 62.8 (d, *J* = 5.3 Hz, CH_2-5'); 43.1 (CH_2-5); 31.2 (CH_2-4); 30.7 (d, *J* = 16.4 Hz, CH_2); 25.4 (d, *J* = 135.1 Hz, CH_2); 20.0 (d, *J* = 4.2 Hz, CH_2). ^{31}P NMR (162 MHz, $D_2O^{25°C}$) *δ* 27.4. Anal Calcd for $C_{16}H_{26}N_3O_9P$: C, 44.14; H, 6.02; N, 9.65; P, 7.11. Found: C, 42.44; H, 5.84; N, 9.37; P, 6.94.

Compound 12. White solid. mp: 134-135 °C. $[\alpha]_D^{26} \cong$ - 8 (*c* 1.0, MeOH). ^{1}H NMR (400 MHz, Methanol-d_4) *δ* 8.00 (d, 1H, *J* = 8.1 Hz, H-6_u); 7.39-7.28 (m, 10H, Ph); 5.95 (d, 1H, *J* = 4.3 Hz, H-1'); 5.78 (d, 1H, *J* = 8.1 Hz, H-5_u); 4.65-4.57 (m, 4H, 4 C*H*$_2$Ph); 4.24-4.18 (m, 3H, H-2', H-3', H-4); 4.14-4.04 (m, 3H, H-4', 2 H-5'); 4.00 (dt, 1H, *J* = 2.6, 1.2 Hz, H-3); 3.57-3.51 (m, 2H, H-2, H-5); 3.41 (dd, 1H, *J* = 12.5, 4.0 Hz, H-5); 1.93-1.84 (m, 2H, 2 H-6); 1.74-1.55 (m, 4H, 2H-7, 2 H-8). ^{13}C NMR (75 MHz, Methanol-d_4) *δ* 166.1 (C); 152.5 (C); 142.6 (CH-6_u); 138.7 (C); 138.6 (C); 129.6 (2 CH_{Ph}); 129.6 (2 CH_{Ph}); 129.2 (2 CH_{Ph}); 129.2 (CH_{Ph}); 129.2 (3 CH_{Ph}); 103.0 (CH-5_u); 90.3 (CH-1'); 85.8 (CH-3); 85.0 (d, *J* = 7.3 Hz, CH-4'); 81.0 (CH-2'); 75.6 (CH-3'); 73.1 (*C*H$_2$Ph); 72.5 (*C*H$_2$Ph); 71.3 (CH-4); 66.5 (CH-2); 64.3 (d, *J* = 5.4 Hz, CH_2-5'); 50.5 (CH_2-5); 33.3 (d, *J* = 13.8 Hz, CH_2); 27.0 (d, *J* = 135.6 Hz, CH_2); 22.1 (d, *J* = 4.4 Hz, CH_2). ^{31}P NMR (162 MHz, Methanol-d_4) *δ* 25.1. Anal Calcd for $C_{30}H_{38}N_3O_{10}P$: C, 57.05; H, 6.06; N, 6.65; P, 4.90. Found: C, 58.31; H, 5.97; N, 6.47; P, 5.02.

Compound 13. White solid. mp: 179-180 °C. $[\alpha]_D^{28} \cong$ - 10 (*c* 1.0, H_2O). ^{1}H NMR (300 MHz, $D_2O^{25°C}$) *δ* 7.96 (d, 1H, *J* = 8.1 Hz, H-6_u); 5.97 (d, 1H *J* = 4.5 Hz, H-1'); 5.95 (d, 1H, *J* = 8.1 Hz, H-5_u);

4.40-4.30 (m, 3H, H-2', H-3', H-4); 4.30-4.25 (m, 1H, H-4'); 4.21-4.04 (m, 3H, H-3, 2 H-5'); 3.55 (dd, 1H, *J* = 12.6, 4.2 Hz, H-5); 3.52-3.43 (m, 1H, H-2); 3.37 (dd, 1H, *J* = 12.6, 2.1 Hz, H-5); 2.07-1.84 (m, 2H, 2 H-6); 1.81-1.60 (m, 4H, 2 H-7, 2 H-8). ^{13}C NMR (75 MHz, $D_2O^{25°C}$) *δ* 166.2 (C); 151.7 (C); 141.6 (CH-6_u); 102.4 (CH-5_u); 88.9 (CH-1'); 83.1 (d, *J* = 7.7 Hz, CH-4'); 78.7 (CH-3); 74.5 (CH-4); 73.8 (CH-2'); 69.5 (CH-3'); 65.4 (CH-2); 62.8 (d, *J* = 5.2 Hz, CH_2-5'); 49.9 (CH_2-5); 31.7 (d, *J* = 16.5 Hz, CH_2); 25.3 (d, *J* = 135.0 Hz, CH_2); 20.2 (d, *J* = 4.2 Hz, CH_2). ^{31}P NMR (162 MHz, $D_2O^{25°C}$) *δ* 27.3. Anal Calcd for $C_{16}H_{26}N_3O_{10}P$: C, 42.58; H, 5.81; N, 9.31; P, 6.86. Found: C, 43.45; H, 5.77; N, 9.03; P, 7.06.

Compound 14. White solid. mp: 119-121 °C. $[\alpha]_D^{22} \cong +13$ (*c* 0.5, MeOH). ^{1}H NMR (400 MHz, Methanol-d_4) *δ* 7.99 (d, 1H, *J* = 8.1 Hz, H-6_u); 7.35-7.23 (m, 15H, Ph); 5.94 (d, 1H, *J* = 4.5 Hz, H-1'); 5.77 (d, 1H, *J* = 8.1 Hz, H-5_u); 4.56-4.47 (m, 6H, 6 CH_2Ph); 4.22-4.16 (m, 2H, H-2', H-3'); 4.11-4.02 (m, 3H, H-4', 2 H-5'); 3.99-3.97 (m, 1H, H-4); 3.87 (dd, 1H, *J* = 4.2, 2.4 Hz, H-3); 3.66-3.55 (m, 3H, H-5, 2 H-9); 3.42-3.34 (m, 1H, H-2); 1.83-1.54 (m, 6H, 2 H-6, 2 H-7, 2 H-8). ^{13}C NMR (75 MHz, Methanol-d_4) *δ* 166.1 (C); 152.5 (C); 142.5 (CH-6_u); 139.1 (C); 139.1 (C); 139.0 (C); 129.5 (2 CH_{Ph}); 129.5 (2 CH_{Ph}); 129.5 (2 CH_{Ph}); 129.3 (2 CH_{Ph}); 129.1 (2 CH_{Ph}); 129.1 (2 CH_{Ph}); 129.0 (CH_{Ph}); 128.9 (2 CH_{Ph}); 103.0 (CH-5_u); 90.1 (CH-1'); 88.2 (CH-3); 85.0 (d, *J* = 7.4 Hz, CH-4'); 84.8 (CH-4); 75.7 (CH-2'); 74.3 (CH_2Ph); 73.0 (CH_2Ph); 72.9 (CH_2Ph); 71.3 (CH-3'); 69.4 (CH_2-9); 64.2 (d, *J* = 5.3 Hz, CH_2-5'); 64.0 (CH-2); 63.7 (CH-5); 34.1 (d, *J* = 14.7 Hz, CH_2); 27.4 (d, *J* = 135.6 Hz, CH_2); 22.0 (d, *J* = 4.2 Hz, CH_2). ^{31}P NMR (162 MHz, Methanol-d_4) *δ* 25.5. Anal Calcd for $C_{38}H_{46}N_3O_{11}P$: C, 60.71; H, 6.17; N, 5.59; P, 4.12. Found: C, 59.76; H, 6.06; N, 5.41; P, 4.23.

Compound 15. White solid. mp: 175-179 °C. $[\alpha]_D^{31} \cong +26$ (*c* 0.9, H_2O). ^{1}H NMR (400 MHz, $D_2O^{25°C}$) *δ* 7.90 (d, 1H, *J* = 8.1 Hz, H-6_u); 5.91 (d, 1H, *J* = 5.0 Hz, H-1'); 5.90 (d, 1H, *J* = 8.1 Hz, H-5_u); 4.32 (t, 1H, *J* = 5.0 Hz, H-2'); 4.28 (t, 1H, *J* = 5.0 Hz, H-3'); 4.24-4.20 (m, 1H, H-4'); 4.11 (ddd, 1H, *J* = 11.9, 4.5, 2.5 Hz, H-5'); 4.03 (ddd, 1H, *J* = 11.9, 5.4, 3.2 Hz, H-5'); 4.01 (t, 1H, *J* = 7.3 Hz, H-4); 3.93 (dd, 1H, *J* = 8.0, 7.3 Hz, H-3); 3.89 (dd, 1H, *J* = 12.7, 3.8 Hz, H-9); 3.82 (dd, 1H, *J* = 12.7, 6.1 Hz, H-9); 3.51 (ddd, *J* = 7.3, 6.1, 3.8 Hz, H-5); 3.38 (dt, *J* = 8.0, 5.6 Hz, H-2); 2.01-1.90 (m, 1H, H-6); 1.89-1.78 (m, 1H, H-6); 1.74-1.58 (m, 4H, 2 H-7, 2H-8). ^{13}C NMR (75 MHz, $D_2O^{25°C}$) *δ* 166.1 (C); 151.7 (C); 141.6 (CH-6_u); 102.4 (CH-5_u); 88.9 (CH-1'); 83.1 (d, *J* = 8.0 Hz, CH-4'); 78.2 (CH-3); 74.3 (CH-4); 73.8 (CH-2'); 69.5 (CH-3'); 62.8 (d, *J* = 5.9 Hz, CH_2-5'); 62.2 (CH-5); 61.0 (CH-2); 58.1 (CH_2-9); 31.3 (d, *J* = 16.8 Hz, CH_2); 25.4 (d, *J* = 135.0 Hz, CH_2); 19.7 (d, *J* = 4.8 Hz, CH_2). ^{31}P NMR (121 MHz, $D_2O^{25°C}$) *δ* 27.5. Anal Calcd for $C_{17}H_{28}N_3O_{11}P$: C, 42.42; H, 5.86; N, 8.73; P, 6.43. Found: C, 41.77; H, 5.81; N, 8.83; P, 6.26.

Compound 16. White solid. mp: 233-236 °C. $[\alpha]_D^{25} \cong +17$ (*c* 0.4, H_2O). ^{1}H NMR (500 MHz, $D_2O^{25°C}$) *δ* 8.01 (d, 1H, *J* = 8.2 Hz, H-6_u); 7.33-7.38 (m, 2H, Ar); 7.01-7.07 (m, 2H, Ar); 6.01 (d, 1H, *J* = 4.3 Hz, H-1'); 5.99 (d, 1H, *J* = 8.2 Hz, H-5_u); 4.41 (dd, 1H, *J* = 4.9, 4.4 Hz, H-2'); 4.38 (dd,1H, *J* = 5.5, 4.7 Hz, H-3'); 4.31-4.35 (m, 1H, H-4'); 4.20-4.26 (m, 1H, H-5'); 4.15 (ddd, 1H, *J* = 8.5, 4.7, 3.1 Hz, H-5'); 3.87 (s, 3H, CH_3O); 3.71 (dd, 1H, *J* = 11.1, 7.8 Hz, H-5); 3.46 (ddd, 1H, *J* = 10.2, 10.1, 4.2 Hz, H-2); 3.19-3.33 (m, 2H, H-4, H-5); 1.67-2.18 (m, 7H, H-3, 2 H-6, 2 H-7, 2 H-8); 1.38-1.47 (m, 2H, $CH_3CH_2CH_2$); 1.07-1.17 (m, 2H, $CH_3CH_2CH_2$); 0.73 (t, *J* = 7.3 Hz, 3H, $CH_3CH_2CH_2$). ^{13}C NMR (125 MHz, $D_2O^{25°C}$) *δ* 166.1 (C); 158.2 (C); 151.7 (C); 141.7 (CH-6_u); 131.2 (C); 129.1 (2 CH_{Ar}); 114.5 (2 CH_{Ar}); 102.5 (CH-5_u); 89.1 (CH-1'); 83.2 (d, *J* = 7.8 Hz, CH-4'); 73.9 (CH-2'); 69.5 (CH-3'); 64.5 (CH-2); 62.9 (d, *J* = 5.3 Hz, CH_2-5'); 55.5 (CH_3O); 50.5 (CH_2-5); 49.7 (CH-4); 48.5 (CH-3); 32.3 (d, *J* = 15.6 Hz, CH_2); 32.0 ($CH_3CH_2CH_2$); 25.7 (d, *J* = 135.1 Hz, CH_2); 20.4 (d, *J* = 4.2 Hz, CH_2); 19.3 ($CH_3CH_2CH_2$); 13.6 ($CH_3CH_2CH_2$). ^{31}P NMR (202 MHz, $D_2O^{25°C}$) *δ* 27.4. Anal Calcd for $C_{26}H_{38}N_3O_9P$: C, 55.02; H, 6.75; N, 7.40; P, 5.46. Found: C, 57.11; H, 6.56; N, 7.42; P, 5.56.

Compound 17. Sticky oil. $[\alpha]_D^{25} \cong -2$ (*c* 1.2, H_2O). ^{1}H NMR (400 MHz; DMSO d_6): δ= 8.23 (s, 1H, C=CH-N); 7.89 (d, 1H, *J*=8.01 Hz, H-6); 7.31 (s, 3H, 3OH); 6.69 (s, 3H, 3OH); 5.79 (d, 1H, *J*=5.66 Hz, H-1'); 5.62 (d, 1H, *J*=8.01Hz, H-5); 5.50 (d, 1H, *J*=9.37 Hz, H-1g); 4.76 (d, 2H, *J*=6.05 Hz, CH_2OP); 4.07 (t, 1H, 5.27 Hz, H-2'); 4.01 (t, 1H, *J*=3.51 Hz, H-3'); 3.94 (t, 1H, *J*=2.15 Hz, H-4'); 3.90-3.80 (m, 2H, H-5'); 3.75 (t, 1H, *J*=9.18 Hz, H-2g); 3.65 (d, 2H, *J*=10.15 Hz, H-6g); 3.46-3.37 (m, 2H, H-3g, H-5g); 3.25-3.18 (m, 1H, H-4g); ^{13}C-NMR (101 MHz; DMSO d_6): δ= 163.27 (C=O); 150.93 (C=O); 144.81 (CH-5_u); 140.96 (=C-N); 123.05 (=CH-N); 102.03 (CH-6_u); 87.50 (CH-1', CH-4'); 83.62 (CH-1g); 79.99 (CH); 77.04 (CH); 73.33 (CH); 72.10 (CH); 70.44 (CH); 69.60 (CH); 64.39 (CH_2); 60.80 (CH_2); 58.00 (CH_2). ^{31}P-NMR (122 MHz; DMSO d_6): δ=-1.79. HRMS Calcd for $[C_{18}H_{26}N_5O_{14}P+H]^+$ 568.1286, Found 568.1288

Compound 18. Sticky oil. $[\alpha]_D^{25} \cong +42$ (*c* 1.6, H_2O). ^{1}H NMR 400 MHz; DMSO-d_6): δ=8.15 (s, 1H, C=CH-N); 7.89 (d, 1H, *J*=8.1 Hz, H-6); 7.30 (br s, 4H, NH_4^+); 6.12 (d, 1H, *J*=5.9 Hz, H-1g); 5.79 (d, 1H, *J*=5.7 Hz, H-1'); 5.61 (d, 1H, *J*=8.1 Hz, H-5); 5.57 (d, 1H, *J*=4.9 Hz, OH); 5.49 (d, 1H, *J*=5.5 Hz, OH); 5.48-5.43 (m, 1H, OH); 5.21 (d, 1H, *J*=4.8 Hz, OH); 5.18 (d, 1H, *J*=5.9 Hz, OH); 4.81-4.72 (m, 2H, CH_2OP); 4.56 (t, 1H, *J*=5.8 Hz, OH); 4.10-4.05 (m, 1H, H-2'); 4.05-4.00 (m, 2H, H-3', H-3g); 3.96-3.92 (m, 1H, H-4'); 3.89-3.78 (m, 2H, H-5_A' e H-5_B'); 3.76-3.69 (m, 1H, H-2g); 3.62 (ddd, 1H, *J*=9.9, 5.3, 2.0 Hz, H-5g); 3.58-3.52 (m, 1H, H-6_Ag); 3.47-3.41 (m, 1H, H-6_Bg); 3.28-3.21 (m, 1H, H-4g). ^{13}C-NMR (101 MHz; DMSO-d_6): δ=163,3 (C=O); 151,0 (C=O); 143,9 (=C-N); 141,0 (CH-5_u); 126,0 (=CH-N); 102.0 (CH-6_u); 87.5 (CH-1'); 85.2 (CH-4'); 83.7 (CH-1g); 76.5 (CH); 73.4 (CH); 72.9 (CH); 70.6 (CH); 70.4 (CH); 69.9 (CH); 64.4 (CH_2); 60.8 (CH_2); 58.0 (CH_2). ^{31}P-NMR (122 MHz; DMSO-d_6): δ=-1.68. HRMS Calcd for $[C_{18}H_{26}N_5O_{14}P+H]^+$ 568.1286, Found 568.1289.

Compound 19. Sticky oil. $[\alpha]_D^{25} \cong -6$ (*c* 2.7, H_2O). ^{1}H NMR (400 MHz; DMSO-d_6): δ=8.08 (s, 1H, C=CH-N); 7.98 (d, 1H, *J*=9.1 Hz, N*H*Ac); 7.90 (d, 1H, *J*=8.1 Hz, H-6); 7.30 (br s, 4H, NH_4^+); 5.79 (d, 1H, *J*=5.8 Hz, H-1g); 5.70 (d, 1H, *J*=9.9 Hz, H-1'); 5.60 (d, 1H, *J*=8.1 Hz, H-5); 5.53-5.30 (m, 4H, 4 OH); 4.82-4.71 (m, 1H, OH); 4.76-4.66 (m, 2H, CH_2OP); 4.09-3.99 (m, 3H, H-2', H-3', H-2g); 3.95-3.91 (m, 1H, H-4'); 3.87-3.76 (m, 2H, H-5_A' e H-5_B'); 3.67 (d, 1H, J=10.8 Hz, H-6_Ag); 3.55 (t, 1H, J=9,2 Hz, H-3g); 3.50-3.36 (m, 2H, H-5g e H-6_Bg); 3.27 (t, 1H, J=9.1 Hz, H-4g); 1.62 (s, 3H, Ac). ^{13}C-NMR (101 MHz; DMSO-d_6): δ=169.5 (C); 163.2 (C=O); 150.9 (C=O); 145.0 (CH-5_u); 141.0 (=C-N); 122.3 (=CH-N); 102.0 (CH-6_u); 87.5 (CH); 86.0 (CH); 83.6 (CH); 80.1 (CH); 74.0 (CH); 73.4 (CH); 70.5 (CH); 70.0 (CH); 64.4 (CH_2); 60.8 (CH_2); 58.0 (CH_2); 54.6 (CH-2g); 22.8 (CH_3). ^{31}P-NMR (122 MHz; DMSO-d_6): δ=-

1.32. HRMS Calcd for $[C_{20}H_{29}N_6O_{14}P+ H]^+$ 609.1552, Found 609.1556

Compound 20. Sticky oil. $[\alpha]_D^{25} \cong +70$ (*c* 2.4, H_2O). ^{1}H NMR (400 MHz; DMSO-d_6): δ=8.16 (s, 1H, C=CH-N); 7.90 (d, 1H, *J*=8.1 Hz, H-6); 7.82 (d, 1H, *J*=6.9 Hz, N*H*Ac); 7.30 (bs, 4H, NH_4^+); 6.22 (d, 1H, *J*=5,8 Hz, H-1g); 5.78 (d, 1H, *J*=5.7 Hz, H-1'); 5.58 (d, 1H, *J*=8.1 Hz, H-5); 4.80-4.71 (m, 2H, CH_2OP); 4.25 (dd, 1H, *J*=10.8, 8.5 Hz, H-3g); 4.10-4.03 (m, 2H, H-2', H-2g); 4.03-3.98 (m, 1H, H-3'); 3.95-3.90 (m, 1H, H-4'); 3.88-3.76 (m, 2H, H-5_A' e H-5_B'); 3.53 (dd, 1H, H=14, 4.7 Hz, H-6_Ag); 3.47-3.40 (m, 2H, H-5g e H-6_Bg); 3.32 (t, 1H, *J*=8.9 Hz, H-4g); 1.67 (s, 3H, Ac). ^{13}C-NMR (101 MHz; DMSO-d_6): δ=170.4 (Ac); 163.2 (C=O); 150.9 (C=O); 144.4 (=C-N); 141.0 (CH-5_u); 125.8 (=CH-N); 102.0 (CH-6_u); 87.5(CH-1'); 83.6 (CH-4'); 83.0 (CH-1g); 76.2 (CH); 73.4 (CH); 73.0 (CH); 70.5 (CH); 70.2 (CH); 64.4 (CH_2); 60.6 (CH_2); 57.9 (CH_2); 53.4 (CH-2g); 22.3 (CH_3). ^{31}P-NMR (122 MHz; DMSO-d_6): δ=-1.35. HRMS Calcd for $[C_{20}H_{29}N_6O_{14}P+ H]^+$ 609.1552, found 609.1556

Acknowledgements

We thank for their support of our programs: MINECO (Madrid, Spain) and FEDER Program (Project CTQ2016-76155-R) and the Government of Aragon (Group E-10) to P.M. A Wellcome Investigator Award (110061) to DMFvA. The Universty of Ferrara (FAR2016) to D.P. and M.F. M.G. thanks MECD for a FPU pre-doctoral grant. D.S. thanks CSIC for a JAE-Pre pre-doctoral grant. KR is funded by a BBSRC Studentship (1416998). Dr. T. Bernardi (University of Ferrara) is acknowledged for ESI-HRMS analysis

Author Contribution

P.M., T.T. and I.D. conceived and planned the experiments. M.G. designed and performed the synthesis of compounds **7-15**. D.S. performed the synthesis of **16**. D.P., N.C, E.M. and M.F. designed and performed the synthesis of compounds **17-20**. P.M. performed docking calculations. I.D. performed molecular dynamic simulations and STD NMR experiments. D.M.F.v.A. and K.R. performed biological assays. P.M. wrote the manuscript with support from D.M.F.v.A., T.T., K.R and I.D.

Keywords: glycosyltransferases • OGT • Nucleotide diphosphate analogues • Carbohydrates • Bioconjugates

[1] (a) J. Voglmeir, S. L. Flitsch, *Sci. Synth., Biocatal. Org. Synth.* **2015**, *1*, 507-542. (b) D.-M. Liang, J.-H. Liu, H. Wu, B.-B. Wang, H.-J. Zhu, J.-J. Qiao, *Chem. Soc. Rev.* **2015**, *44*, 8350-8374. (c) M. M. Palcic, *Curr. Opin. Chem. Biol.* **2011**, *15*, 226-233.

[2] S. Kellokumpu, A. Hassinen, T. Glumoff, *Cell. Mol. Life Sci.* **2016**, *73*, 305-325.

[3] (a) D. Horton, in *Carbohydrate Chemistry, Biology and Medical Applications* (Eds.: M. K. Cowman, C. A. Hales), Elsevier, Oxford, **2008**, pp. 1-28. (b) R. W. Gantt, P. Peltier-Pain, W. J. Cournoyer, J. S. Thorson, *Nat. Chem. Biol.* **2011**, *7*, 685-691. (c) C. Zhang, B. R. Griffith, Q. Fu, C. Albermann, X. Fu, I.-K. Lee, L. Li, J. S. Thorson, *Science* **2006**, *313*, 1291-1294.

[4] M. E. Etzler, A. Varki, R. L. Cummings, J.D. Esko, H. H. Freeze, G. W. Hart, *Essentials of Glycobiology, 2nd Ed.* Cold Spring Harbor Laboratory Press, New York, **2008**.

[5] (a) Z. Ma, K. Vosseller, *Amino Acids* **2013**, *45*, 719-733. (b) T. P. Lynch, M. J. Reginato, *Cell Cycle* **2011**, *10*, 1712-1713.

[6] (a) Y. Zhu, X. Shan, S. A. Yuzwa, D. J. Vocadlo, *J. Biol. Chem.* **2014**, *289*, 34472-34481. (b) S. A. Yuzwa, D. J. Vocadlo, *Chem. Soc. Rev.* **2014**, *43*, 6839-6858.

[7] X.-L. Sun, *Med. Chem.* **2013**, *3*, e106.

[8] (a) S. Wang, S. Vidal, *Carbohydr. Chem.* **2013**, *39*, 78-101. (b) R. Roychoudhury, N. L. B. Pohl, *Curr. Opin. Chem. Biol.* **2010**, *14*, 168-173. (c) J. Schutzbach, I. Brockhausen, *Methods in Molecular Biology (Totowa, NJ, United States)* **2009**, *534*, 359-373. (d) E. Walker-Nasir, I. Ahmad, M. Saleem, D. C. Hoessli, *Curr. Org. Chem.* **2007**, *11*, 591-607. (e) K.-H. Jung, R. R. Schmidt, *Carbohydr.-Based Drug Discov.* **2003**, *2*, 609-659.

[9] T. Kajimoto, M. Node, *Synthesis* **2009**, 3179-3210.

[10] (a) X. Qian, M. M. Palcic, *Carbohydr. Chem. Biol.* **2000**, *3*, 293-312. (b) L. J. Whalen, W. A. Greenberg, M. L. Mitchell, C.-H. Wong, *Iminosugars* **2007**, 153-176. (c) L. Cipolla, B. La Ferla, M. Gregori, *Comb. Chem. High Throughput Screening* **2006**, *9*, 571-582. (d) P. Compain, O. R. Martin, *Bioorg. Med. Chem.* **2001**, *9*, 3077-3092.

[11] P. Merino, I. Delso, T. Tejero, M. Ghirardello, V. Juste-Navarro, *Asian J. Org. Chem.* **2016**, *5*, 1413-1427

[12] (a) L. Tedaldi, G. K. Wagner, *MedChemComm* **2014**, *5*, 1106-1125. (b) J. Jiang, M. B. Lazarus, L. Pasquina, P. Sliz, S. Walker, *Nat. Chem. Biol.* **2012**, *8*, 72-77.

[13] (a) D. A. Smithen, S. M. Forget, N. E. McCormick, R. T. Syvitski, D. L. Jakeman, *Org. Biomol. Chem.* **2015**, *13*, 3347-3350. (b) M. Izumi, H. Yuasa, H. Hashimoto, *Curr. Top. Med. Chem.* **2009**, *9*, 87-105.

[14] (a) S. Wang, J. A. Cuesta-Seijo, D. Lafont, M. M. Palcic, S. Vidal, *Chem. Eur. J.* **2013**, *19*, 15346-15357. (b) S. Wang, J. A. Cuesta-Seijo, A. Striebeck, D. Lafont, M. M. Palcic, S. Vidal, *ChemPlusChem* **2015**, *80*, 1525-1532.

[15] M. Ghirardello, M. de las Rivas, A. Lacetera, I. Delso, E. Lira-Navarrete, T. Tejero, S. Martin-Santamaria, R. Hurtado-Guerrero, P. Merino, *Chem. Eur. J.* **2016**, *22*, 7215-7224.

[16] (a) E. Lira-Navarrete, M. d. l. Rivas, I. Companon, M. C. Pallares, Y. Kong, J. Iglesias-Fernandez, G. a. J. L. Bernardes, J. M. Peregrina, C. Rovira, P. Bernado, P. Bruscolini, H. Clausen, A. Lostao, F. Corzana, R. Hurtado-Guerrero, *Nat. Commun.* **2015**, *6*, 6937. (b) R. Hurtado-Guerrero, *Biochem. Soc. Trans.* **2016**, *44*, 61-67. (c) D. Madariaga, N. Martinez-Saez, V. J. Somovilla, L. Garcia-Garcia, M. Alvaro Berbis, J. Valero-Gonzalez, S. Martin-Santamaria, R. Hurtado-Guerrero, J. L. Asensio, J. Jimenez-Barbero, A. Avenoza, J. H. Busto, F. Corzana, J. M. Peregrina, *Chem. Eur. J.* **2014**, *20*, 12616-12627.

[17] (a) S. Pathak, J. Alonso, M. Schimpl, K. Rafie, D. E. Blair, V. S. Borodkin, A. W. Schuttelkopf, O. Albarbarawi, D. M. F. van Aalten, *Nat. Struct. Mol. Biol.* **2015**, *22*, 744-750. (b) Y. Zhu, T.-W. Liu, S. Cecioni, R. Eskandari, W. F. Zandberg, D. J. Vocadlo, *Nat. Chem. Biol.* **2015**, *11*, 319-325. (c) M. B. Lazarus, J. Jiang, T. M. Gloster, W. F. Zandberg, G. E. Whitworth, D. J. Vocadlo, S. Walker, *Nat. Chem. Biol.* **2012**, *8*, 966-968. (d) A. Ostrowski, D. M. F. van Aalten, *Biochem. J.* **2013**, *456*, 1-12.

[18] T. M. Gloster, D. J. Vocadlo, *Curr. Signal Transduct. Ther.* **2010**, *5*, 74-91.

[19] (a) B. J. Gross, B. C. Kraybill, S. Walker, *J. Am. Chem. Soc.* **2005**, *127*, 14588-14589. (b) H. C. Dorfmueller, V. S. Borodkin, D. E. Blair, S. Pathak, I. Navratilova, D. M. F. Aalten, *Amino Acids* **2011**, *40*, 781-792.

[20] (a) S. Wang, D. L. Shen, D. Lafont, A.-S. Vercoutter-Edouart, M. Mortuaire, Y. Shi, O. Maniti, A. Girard-Egrot, T. Lefebvre, B. M. Pinto, D. Vocadlo, S. Vidal, *MedChemComm* **2014**, *5*, 1172-1178. (b) A. K. Nagel, L. E. Ball, *Amino Acids* **2014**, *46*, 2305-2316.

[21] R. Trapannone, K. Rafie, D. M. F. van Aalten, *Biochem. Soc. Trans.* **2016**, *44*, 88-93.

[22] (a) M. Mitchell, L. Qaio, C.-H. Wong, *Adv. Synth. Catal.* **2001**, *343*, 596-599. (b) M. D. Burkart, S. P. Vincent, A. Duffels, B. W. Murray, S. V. Ley, C. H. Wong, *Bioorg. Med. Chem.* **2000**, *8*, 1937-1946. (c) P. Merino, I. Delso, T. Tejero, M. Ghirardello, V. Juste-Navarro, *Asian J. Org. Chem.* **2016**. (d) P. Merino, T. Tejero, I. Delso, R. Hurtado-Guerrero, A. Gomez-SanJuan, D. Sadaba, *Mini-Rev. Med. Chem.* **2012**, *12*, 1455-1464.

[23] A. Ardevol, C. Rovira, *J. Am. Chem. Soc.* **2015**, *137*, 7528-7547.

[24] K. K. Yeoh, T. D. Butters, B. L. Wilkinson, A. J. Fairbanks, *Carbohydr. Res.* **2009**, *344*, 586-591.

[25] I. Delso, T. Tejero, A. Goti, P. Merino, *Tetrahedron* **2010**, *66*, 1220-1227.

[26] P. Merino, I. Delso, V. Mannucci, T. Tejero, *Tetrahedron Lett.* **2006**, *47*, 3311-3314.

[27] A. Dondoni, S. Staderini, A. Marra, *Eur. J. Org. Chem.* **2013**, 5370-5375.

[28] (a) M. G. Finn, V. V. Fokin, *Chem. Soc. Rev.* **2010**, *39*, 1231-1232. (b) H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem., Int. Ed.* **2001**, *40*, 2004-2021. (c) C. Wang, D. Ikhlef, S. Kahlal, J.-Y. Saillard, D. Astruc, *Coord. Chem. Rev.* **2016**, *316*, 1-20. (d) L. Liang, D. Astruc, *Coord. Chem. Rev.* **2011**, *255*, 2933-2945.

[29] V. K. Tiwari, B. B. Mishra, K. B. Mishra, N. Mishra, A. S. Singh, X. Chen, *Chem. Rev.* **2016**, *116*, 3086-3240.

[30] (a) S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* **1981**, *22*, 1859-1862. (b) M. H. Caruthers, S. L. Beaucage, C. Becker, J. W. Efcavitch, E.

F. Fisher, G. Galluppi, R. Goldman, P. DeHaseth, M. Matteucci, et al. *Gene Amplif. Anal.* **1983**, *3*, 1-26. (c) S. L. Beaucage, *Methods Mol. Biol. (Totowa, N. J.)* **1993**, *20*, 33-61.

[31] K. Rafie, A, Gorelik, R. Trapannone, V. Borodking, D. M. F. van Aalten, submitted manuscript. (submitted manuscript). For details see SI).

[32] M. Schimpl, X. Zheng, V. S. Borodkin, D. E. Blair, A. T. Ferenbach, A. W. Schüttelkopf, I. Navratilova, T. Aristotelous, O. Albarbarawi, D. A. Robinson, M. A. Macnaughtan, D. M. F. v. Aalten, *Nat. Chem. Biol.* **2012**, *8*, 969-974.

[33] M. Mayer, B. Meyer, *Angew. Chem. Int. Ed.* **1999**, *38*, 1784-1788.

[34] M. d. C. Fernández-Alonso, D. Díaz, M. Á. Berbis, F. Marcelo, J. Cañada, J. Jiménez-Barbero, *Curr. Protein Peptide Sci.* **2012**, *13*, 816-830.

[35] T. Haselhorst, A. C. Lamerz, M. v. Itzstein, *Methods Mol. Biol.* **2009**, *534*, 375-386.

[36] X. Ma, P. Liu, H. Yan, H. Sun, X. Liu, F. Zhou, L. Li, Y. Chen, M. M. Muthana, X. Chen, P. G. Wang, L. Zhang, *PLoS One* **2013**, *8*, e63452.

[37] M. B. Lazarus, Y. Nam, J. Jiang, P. Sliz, S. Walker, *Nature* **2011**, *469*, 564-567.

[38] M. Kumari, S. Kozmon, P. Kulhanek, J. Stepan, I. Tvaroska, J. Koca, *J. Phys. Chem. B* **2015**, *119*, 4371-4381.

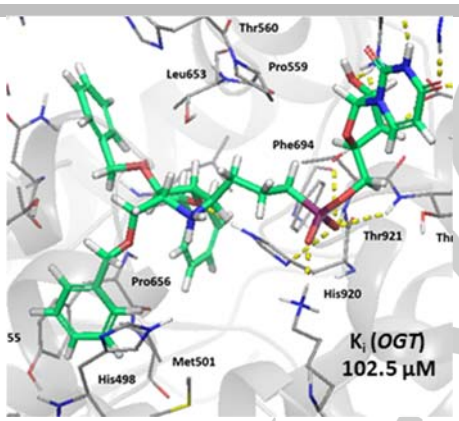
[39] S. Pathak, J. Alonso, M. Schimpl, K. Rafie, D. E. Blair, V. S. Borodkin, A. W. Schuttelkopf, O. Albarbarawi, D. M. F. van Aalten, *Nat. Struct. Mol. Biol.* **2015**, *22*, 744-750.

Entry for the Table of Contents (Please choose one layout)

Layout 1:

FULL PAPER

The lower affinity caused by the absence of β-phosphate in UDP-GlNAc analogues is offset by the presence of hydrophobic groups at the pyrrolidine used as a surrogate of the carbohydrate unit



*M. Ghirardello, D. Perrone, N. Chinaglia, D. Sádaba, I. Delso, T. Tejero, E. Marchesi, M. Fogagnolo, K. Rafie, D. M. F. van Aalten, and P. Merino**

Page No. – Page No.

UDP-GlcNAc Analogs as Inhibitors of *O*-GlcNAc Transferase: Spectroscopic, Computational and Biological Studies